**DA 09-2433**
**Released: November 18, 2009**
**Comment Date: December 9, 2009**

**COMMENT SOUGHT ON DATA PORTABILITY AND ITS RELATIONSHIP TO BROADBAND**
**NBP Public Notice #21**

**GN Docket Nos. 09-47, 09-51, and 09-137**

Comments of the InCommon Steering Committee – NBP Public Notice #21.

Comments are provided by the InCommon Steering Committee, a formal group of leading CIOs and IT leaders within the higher education community and is the governing body of the InCommon Federation. InCommon, a project of Internet2, is the U.S. identity and access management of higher education and its partners. The mission of the InCommon Federation is to create and support a common framework for trustworthy shared management of access to on-line resources in support of education and research in the United States. To achieve its mission, InCommon facilitates development of a community-based common trust fabric sufficient to enable participants to make appropriate decisions about the release of identity information and the control of access to protected online resources. InCommon is intended to enable production-level end-user access to a wide variety of protected resources.

1. **Government data transparency.** Data transparency refers to making data public and easily accessible over the Internet. There are many pieces of legislation requiring the publication of Federal government information.[1] This legislation typically requires the publication of data on an agency's website. One recent initiative seeks to establish a central repository of government data.[2] We seek comment on the potential benefits and pitfalls of increased data transparency.
    a. What efficiencies can be gained through easing accessibility to public government information?

**Internet communication is how people access information: locally, nationally and globally. Efficiencies are built into that process. The benefit of having a central database would be to ensure that all information is collected and posted equally rendering appropriate public access. One concern, however, is the potential to violate federal privacy laws that prohibit the intermingling of data collected and maintained by the federal government. Promulgated almost forty years ago, in light of new technologies, altered social norms and needs, these laws need revising to maintain their intent that data not be misused against an individual or groups of individuals (suspect classes in particular) while allowing for new formats and access accountability.**

b. Are there examples of innovative products or services provided by the private sector that rely upon the use of easily accessible government information?

**No comment**

[1] *See* Federal Funding Accountability and Transparency Act of 2006, Pub. L. No. 109-282, 120 Stat. 1186 (2006); *Recovery Act*

[2] *See* Data.gov (www.Data.gov last visited November 18, 2009).

c. Federal government data are available in many formats. In what formats should this data be made available over the Internet? How should open data standards inform policy for data transparency?

**Open standards must be the foundational technological format for federal government sites in keeping with the alignment that already exists by definition between the not for profit nature of such formats and the services the government offers to its people. It would not be an exaggeration to state that open formats as a technological business model reflect the fundamental principles of democracy as a political system.**

d. How does data transparency relate to application development? Are there potential efficiencies to be gained through an increase in government data transparency?

**Data transparency and application development are mutually reinforcing technologies. Following Sir Tim Berners-Lee's original vision of a machine-readable Web, the availability of data creates vast opportunities for new application development. Conversely, hiding data behind application programming interfaces (APIs) uses scarce government programming resources needed to address the widely varied needs of researchers. Serendipitous and direct access to data fosters ideas for new applications without prior negotiation of API features and related expenses by both data providers and application developers.**

e. To what extent would increased data transparency affect intra-agency processes, intergovernmental coordination, and civic participation?

**If done properly this effort should have a meaningful, salutary effect on intra-agency process, intergovernmental coordination and civil participation. As noted above under section 1(a), the first concern should be the collision with existing privacy laws that prohibit such activity while maintaining appropriate technical safeguards and privacy practices; if this effort requires a change in the laws that fact, too, is beneficial because privacy law reform has long been on the horizon as an important area of legal reform in this country. The second concern would be to map out the effort comprehensively. In so doing, more natural inter-agency connections may emerge than exist in the structure today. For one example, an agency devoted to Internet related issues could, conceivably, emerge from the process, instead of this powerful new technology straining the boundaries of multiple other agencies such as the Copyright Office, FCC, FTC, the Judiciary, Department of Commerce, Department of Defense, and the Department of Homeland Security. A less stultified, more dynamic federal government responsive to the participatory movement among users of the Internet would, once again, complement and reflect the underlying principles of U.S. democracy.**

f. To what extent do existing regulations inhibit or promote government data transparency?

**The Privacy Act of 1974 is the most obvious example of potential collision with new technologies that provide access to government information. The Electronic Communications Privacy Act is sorely in need of updating since in its existing form, collapsing telephony and data networking technologies, it creates conflict with existing Fourth Amendment jurisprudence; specifically, the difference between the technologies does not map to the existing rubric between the processes necessary for the divide between the release of content and non-content types of information to law enforcement, if not others. Absences in law need stating to make the portrait of this effort more complete. Two**

**examples are oversight rules to manage the use the federal government makes of third party vendors to gather information in criminal investigations and a federal data breach notification law that harmonizes existing state rules. The most important absence, however, goes to the heart of the question regarding this effort towards greater governmental transparency: comprehensive privacy laws especially regarding data.**

g. What impact do developments in data transparency have with respect to broadband deployment, adoption, and use?

**Broadband represents the physical layer necessary to effectuate full access to government information. The relatively low rating of the U.S. on broadband deployment is well known, although sometimes the comparison would benefit from appreciating the geographic expanse by comparison to more highly ranked countries such as South Korea. This geographic expanse represents a special though not insurmountable challenge. A variety of methods, including a tax on Internet connections not unlike the tax on telephone connections that the federal government used to bring electricity to rural areas where it was not cost-efficient for corporations, could be deployed to achieve this goal. Finally, the Internet is made up of three essential layers: physical, logical and application. All three layers are necessary to realize the potential of transparency.**

h. What are the potential benefits to making data more accessible?

**There is considerable danger to the dignity of the individual, which is the cornerstone of healthy democracy, of blithely pursuing transparency without a comprehensive and current updating of U.S. privacy laws regarding data about or affecting the individual, especially since the power that new technologies have to gather and manipulate data can and are already being used deleteriously by corporations, governments and criminals.**

i. What potential pitfalls exist when increasing data transparency?

**The failure to map the effort out comprehensively and with an eye not merely to technology but to law, business models, social norms and needs could create serious adverse effects of unintended consequences as wide reaching as from national security to, as noted above, an undermining of the rights and dignity of the individual.**

j. What privacy and confidentiality concerns might arise due to an increase in data transparency and what, if any, privacy safeguards are needed to protect against the misuse of personal information?

**A number of thoughtful scholars have detailed in the last decade the real and potential abuses that governments, corporations and criminals make of information about the individual. Here we speak primarily of the work of Daniel Solove, although many others could be named. Any effort to map out this program must incorporate the insights of thinkers who study the psychological, social and political dimensions of privacy.**

k. What types of personal information should be protected from disclosure?

**The usual list will be obvious for many such as social security or national identification numbers, bank routing, credit cards, driver's license and perhaps date of birth, internet protocol addressing or any**

**other technological routing and/or device registration mechanisms, in addition to the already protected health information contained in patient treatment records. What requires stating at this juncture is that governments, corporations and criminals will shift this target quickly and that management of such an effort requires an on-going oversight with sufficient powers (and checks) to react nimbly to new threats of disclosed and/or recombined data as well as strong technological safeguards against automated collection.**

2. **Cloud computing.** When considering the portability of data, we also consider the processes through which data are moved. In this context, we seek comment on how to identify and understand cloud computing as a model for technology provisioning.
    a. The National Institute of Standards and Technology defines cloud computing as "a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction."[3] Does this definition accurately capture the concept of cloud computing?

**Yes, because it does not restrict the definition to for-profit corporations for provisioning.**

b. What types of cloud computing exist (e.g., public, hybrid, and internal) and what are the legal and regulatory implications of their use?

**Thought at the federal level about the implications of cloud computing is necessary and may come to include regulatory activity. For example, the export of data to foreign locations that, if breached or stolen, could result in a threat to national security on a broad scale beginning with military activity and extending to economic empowerment is a notable problem. At home, regulatory safeguards should be instituted to be sure that individuals and institutions that use corporate services are sufficiently protected from a misappropriation of their data in the event of breach, theft or shifts in corporate ownership. Several issues that at the moment are the focus of contract relations between individuals and entities, such as higher education institutions, might conceivably become part of a package of regulated areas such as indemnity, warranty, export control issues, forum location for legal suit or conflict resolution, responsibility for breach notification, data mining, user tracking, and consumer protection, just to name a few.**

c. Can present broadband network configurations handle a large-scale shift in bandwidth usage that a rapid adoption of cloud computing might cause?

**Present broadband configurations evolved during the growth of conventional services such as web browsing and video delivery. Rapid change that occurs because of an entirely new type of service could seriously decrease the quality of service delivery broadly. Since it takes time to reconfigure or upgrade infrastructure, this issue should be studied and understood well ahead of such a large-scale shift.**

d. How does cloud computing affect the reliability, scalability, security, and sustainability of information and data?

[3] *See* CDT Comments at 13.

**Obvious selling points at this juncture are scalability and to some degree reliability but as noted in 2(b) above, security, privacy (as a matter separate from technological safeguards), and even sustainability are in question without either or both improved bargaining relations between entities that provision services and individuals and entities as clients.**

e. To what extent can the federal government leverage cloud solutions to improve intra-agency processes, intergovernmental coordination, and civic participation?

**As a technology, cloud computing presents some challenges, bandwidth among them, but otherwise is virtually unlimited. To date, its market restraints lie largely in the confusion surrounding responsibility for and ownership of the information transmitted and stored. In order to facilitate business models in keeping with the technological promise of cloud computing, the government should therefore consider regulation to harmonize the privacy and security requirements for all entities that store or transmit sensitive information relative to individual persons. The requirements would be uniform among governments, cloud computing service providers and entities that obtain personally identifiable or other forms of sensitive data based on their relationship to the individual, as well as third parties that, as an industry type and matter of practice, mine, collect and collate information about individuals. This effort would begin necessarily with a sufficiently broad definition to incorporate records currently protected by law, such as education and medical records, but would include any and all data elements that in combination result in the identification of an individual and disclosure of information about that individual that would have a deleterious effect on his or her autonomy to function unhindered in society (for example, without cause, to obtain employment, a mortgage or loan, purchase or sell goods or services, gain admission to an educational institution, practice a chosen faith and speak or associate and identify one's self freely in the culture's context.) The next step would be to establish baseline administrative, logical and physical security requirements for that data irrespective of the particular form they are stored or transmitted. Fundamental fair information practices should be applied to that information, namely transparency about what information is collected, for what purpose, how long it is maintained, and the rules adhering to disclosure (to whom, for what reason and whether or not notice is required). Finally, uniform rules regarding information breach notification should be applied including the standard definition of a breach that includes criteria by which that determination is made (for example, for an electronic breach, criteria such as the nature of the vulnerability, the type of tool used to breach technical security, amount of data transmitted and for how long the breach occurred), the elements of notice and any permissible defenses (for example, encryption of data).**

f. What impact do developments in cloud computing have with respect to broadband deployment, adoption, and use?

**Broadband supplies the necessary physical layer to access these services.**

g. How can various parties leverage cloud computing to obtain economic or social efficiencies? Is it possible to quantify the efficiencies gained?

**No comment**

h. To what extent are consumers protected by industry self-regulation (e.g., the Cloud Computing Manifesto[4]), and to what extent might additional protections be needed?

**See answer to 2(b) above**

i. What specific privacy concerns are there with user data and cloud computing?

**See answer to 2(b) above, in conjunction with answers to subsections under 1 above related to privacy.**

j. What precautions should government agencies take to prevent disclosure of personal information when providing data?

**See answer to 2(b) above, in conjunction with answers to subsections under 1 above related to privacy.**

k. Is the use of cloud computing a net positive to the environment? Are there specific studies that quantify the environmental impact of cloud computing?

**No comment**

3. **Identity Management and Government Service Delivery.**[5] Data held by the government may be personally sensitive or confidential. In this context, we seek comment on identity management as it relates to the provision of services where individuals either provide data to the government or access data that are personally sensitive or confidential.
   a. What is the current state of identity management in the federal, state, local and Tribal government?

**See Federal Identity, Credential, and Access Management (FICAM) Roadmap and Implementation Guidance Version 1.0, November 10, 2009, http://www.idmanagement.gov/documents/FICAM_Roadmap_Implementation_Guidance.pdf**

b. What is the spectrum of online identity credentialing required for access to online services from the government and non-governmental entities?

**At present the spectrum is very narrow and constrained by limitations of current identity management systems. Both management of appropriate access and appropriate management of access will require a richer notion of identity that includes such things as affiliations, roles and responsibilities. For example, access in emergency response situations requires credentialing of authority and responsibility is essential in addition to a unique identifier. In other use cases, individual identifiers may not be needed or appropriate and their provision could violate a user's privacy. Richer identity elements will be garnered from multiple authoritative sources, and must include strong protections of individual privacy for reasons of both security and individual freedom.**

[4] Cloud Computing Manifesto, http://wiki.cloudcommunity.org/wiki/Cloud_Computing_Manifesto (last visited Aug. 20, 2009).

[5] By identity management, we refer to electronic identity credentialing. See IDManagement.gov for more details.

**Most government services can be accessed by a password. However, some services, such as grant applications and research administration are beginning to require a form of two-factor identification. See Federal Identity, Credential, and Access Management (FICAM) Roadmap and Implementation Guidance Version 1.0, November 10, 2009, http://www.idmanagement.gov/documents/FICAM_Roadmap_Implementation_Guidance.pdf**

**Non-governmental use encompasses all of the products and services available.**
**Broadband communications services are meant to move very large amounts of information very quickly. The results of this movement, while beneficial in most cases, also can have serious adverse consequences if done by misfeasance or malfeasance. As such, the management of access to information as well as to control of the communication infrastructure must be robust and as rapid as the infrastructure itself.**

**One of the weak areas today is the Internet domain name translation system itself. Manipulation of that system can divert information flows to effect theft or denial of service. Of even greater concern is the next generation infrastructure, which will allow manipulation of the substructure carrying these very high speed information flows. While this is an important new type of broadband service, management of access to these capabilities must be extremely robust.**

**In light of the above, identity management of both individual sources of command and control (users and platforms) as well as infrastructure components is critical. All infrastructure components should be built on trusted computing platforms and software layer components should have individual identities. Assertion and verification of identities and permissions must be as rapid as data flows, which implies significant advances in cryptographic processing power. Research into how to achieve this level of trusted infrastructure must precede widespread adoption of next generation broadband technologies.**

c. What identity management technologies currently exist and what are their applications?

**See Federal Identity, Credential, and Access Management (FICAM) Roadmap and Implementation Guidance Version 1.0, November 10, 2009, http://www.idmanagement.gov/documents/FICAM_Roadmap_Implementation_Guidance.pdf. One specific technology in this document is federated identity management such as InCommon for higher education. The power of federated identity is that each identity provider in the federation makes its own decisions about how it will credential the individuals associated with it, while each service provider decides whether or not the standards used by each identity provider meet that service providers requirements. Each party, identity provider and service provider, thus maintains control of authentication and authorization respectively. Trust, privacy, and access are mediated through the proper allocation of controlled policies and procedures agreed upon by each entity.**

d. How have HSPD-12 implementation efforts affected the efficiency of the federal government?[6]

[6] *See* Homeland Security Presidential Directive-12: Policy for a Common Identification Standard for Federal Employees and Contractors (http://www.dhs.gov/xabout/laws/gc_1217616624097.shtm last visited November 18, 2009).

**No comment**

e. What identity management technologies are available in the private sector? What are their applications?

**See Federal Identity, Credential, and Access Management (FICAM) Roadmap and Implementation Guidance Version 1.0, November 10, 2009, http://www.idmanagement.gov/documents/FICAM_Roadmap_Implementation_Guidance.pdf**

f. What impact do developments in identity management, such as Open ID,[7] have with respect to broadband deployment, adoption, and use?

**Such developments should have no impact per se on broadband deployment, adoption, and use. OpenID might have security weaknesses. Therefore, it is the responsibility of the broadband service provider to perform a risk analysis to decide if OpenID, or any other candidate identity protocol, is strong enough to be used to access that provider's service.**

g. What are the potential benefits of a coordinated nationwide identity management schema?

**The benefits of a coordinated identity management system are that information can be transferred among the various service providers while protecting the identity of those who access them. This provides interoperability that means better service for customers. In addition to interoperability, an appropriately deployed identity management system can become woven into the trust fabric of an individual's relationship with the mediating entity, whether that is a higher education institution, a private corporation or the government. To the degree that a trusted identity management system can be sufficiently coordinated in terms of both technology and policy, managed appropriately, it would ease an individual's use of the Internet for any variety of public and private purposes while enhancing privacy and security. The equation requires all three components: usability, privacy and security.**

h. What are the potential pitfalls of a coordinated nationwide identity management strategy?

**While other countries provide a national identity that can be used to access government provided services, identity management is not yet mature as a technology to know that a single identifier is all that is needed for all services. Nor should it be. A single identifier, if compromised, would seriously harm the actual subject, perhaps permanently. An individual should be able to acquire multiple abstract identifiers when necessary. Such identifiers could be discarded and new ones issued without affecting the valid use of the others. Possibly, different communities of interest may have reasons for requiring different identities.**

**Moreover, some scholars have postulated that identity management supports an unwanted proprietary use of information, and provides, in the case of breach or theft, a trove of information that can be misused at the expense of the entity standing up the network or system as well as to the individuals whose personal identities have been compromised. To the first point, it goes to other**

[7] *See* IDManagement.gov. (http://idmanagement.gov/ last visited November 18, 2009).

**questions regarding the economies of intellectual property, which are separate from this inquiry; to the second point, these risks, while weighty, are nonetheless understood and urge greater emphasis on research, promulgation and implementation of robust privacy laws, regulations and practices as well as ever-current technological safeguards.**

i. What specific privacy concerns are there with identity management strategies?

**Technical security is the most specific privacy concern related to identity management strategies. Since technical security is "only as strong as its weakest link" it is important that an identity management strategy be comprehensive and holistic taking into account the physical, logical and application layers upon which it operates and communicates with other entities and resource providers. Policy plays a critical role in negotiating privacy in this space by establishing the minimum standards upon which a partnership is based, the data networking environment of that communication and the expectations regarding both metadata and content exchanged in the relationship.**

j. What types of personal information should be protected from disclosure?

**All personally identifiable information must be protected from disclosure. While it may be that a single element of personal information may not permit harm to be done to an individual, certain single items, for example, the social security number may make it possible to determine the name (or vice versa) and with name and social security number much damage may be done. Generally, the disclosure of any two personally identifiable pieces of information regarding an individual is sufficient to permit harm to that person. And as information technologies become more powerful, it is increasingly possible to mine smaller and, in isolation more esoteric, attributes of an individual and recombine them in ways that constitute the identity of a specific individual. Although a number of studies have already demonstrated these features of information technologies, for example see the work of Hal Abelson, Ken Ledeen and Harry Lewis, the full potential of encroachment on the cultural meaning of individual dignity, from economic security to emotional well-being, is not yet understood clearly. Therefore it is critical that any efforts that the government makes towards greater transparency, or that which involves any form of particularized information about individuals, must be made with a full appreciation of potential unintended consequences, subject to nimble and immediate remediation and open to constant adjustment and change as the needs of a democratic society requires.**

Respectfully submitted,

Gary Bachula/[illegible]

**The InCommon Steering Committee**

Gary Bachula
Vice President for External Relations
Internet2
1150 18th Street, NW
Suite 1020
Washington, DC 20036